Δείκτες αναγνωσιμότητας: Ζητήματα εφαρμογής και αξιοπιστίας

Δημήτριος Τζιμώκας¹ και Μαρίνα Ματθαιουδάκη²

¹Κέντρο Ελληνικής Γλώσσας, ²Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
dtzimok@phil.uoa.gr, marmat@enl.auth.gr

Abstract

This paper presents some introductory information about the notion of readability, the development of some widely known readability formulas for both English and Greek and the online readability tool “grval 1.1”, which was developed by researchers at the Centre for the Greek Language. The paper aims to test the consistency of ‘grval 1.1’ and to this aim the readability measures of “grval 1.1” on a sample of reading material were measured. The material was taken from the “Reading comprehension” module which is part of the examinations for attainment in Greek of the Centre for the Greek Language (CGL). The sample included reading material for all four levels of attainment in Greek used in the exam papers between the years 2007 and 2010. The analysis indicated low consistency of the readability tool and proposals for addressing the inconsistencies are made.

1. Εισαγωγή

1.1. Προσδιορισμός της έννοιας της αναγνωσιμότητας

Μεταξύ των διαφόρων ορισμών της έννοιας της *αναγνωσιμότητας* (readability) που παραθέτει ο DuBay (2004: 3), ο πιο κατανοητός είναι των Edgar Dale και Jeanne Chall (1948). Οι ερευνητές αυτοί υποστηρίζουν ότι η αναγνωσιμότητα συνιστά το συνολικό άθροισμα όλων εκείνων των χαρακτηριστικών (συμπεριλαμβανομένων των μεταξύ τους αλληλεπιδράσεων) ενός συγκεκριμένου δείγματος έντυπου υλικού, τα οποία επηρεάζουν τον βαθμό επιτυχούς ανάγνωσής του από μια ομάδα αναγνωστών. Ο βαθμός επιτυχούς ανάγνωσης ενός κειμένου έγκειται στο βαθμό που οι αναγνώστες το κατανοούν, διατηρούν την προσλαμβανόμενη από αυτό πληροφορία, το διαβάζουν με ικανοποιητική ταχύτητα και το βρίσκουν ενδιαφέρον. Αναμφίβολα, λοιπόν, η αναγνωσιμότητα ή αλλιώς αναγνωστική δυσκολία είναι το αποτέλεσμα της διάδρασης μεταξύ κειμένου και αναγνώστη. Τα χαρακτηριστικά των δύο αυτών κύριων συνιστωσών της αναγνωσιμότητας είναι τα εξής: αφενός το κείμενο συμβάλλει με το περιεχόμενό του, το ύφος του, το

σχεδιασμό του (μέσο, φωτογραφίες, γραμματοσειρά, χρώμα κτλ.) και τη δομή του· αφετέρου ο αναγνώστης συμβάλλει με την πρότερη γνώση του, την αναγνωστική του δεξιότητα, το ενδιαφέρον του και το κίνητρό του (βλ. Gray & Leary 1935, Γιάγκου 2009: 152-154). Εάν βάλουμε κατά μέρος τα χαρακτηριστικά του αναγνώστη, τα οποία είναι εξ ορισμού δυσερευνήτα, από τα χαρακτηριστικά του κειμένου που μας απομένουν είναι προφανές ότι μηχανικά μετρήσιμο και επομένως το μόνο ικανό να δώσει μια κατά το δυνατόν αντικειμενική πρόβλεψη για το βαθμό επιτυχούς ανάγνωσης ενός κειμένου είναι το στοιχείο του ύφους. Με τον όρο ‘ύφος’ εννοείται το σύνολο των στοιχείων που σχετίζονται με τη λεξικοποίηση σε ένα κείμενο των εννοιών που αυτό πραγματεύεται και οι μεταξύ τους σχέσεις. Πρόκειται για τα λεγόμενα επιφανειακά χαρακτηριστικά ενός κειμένου, δηλαδή το λεξιλόγιο, τη δομή των προτάσεων και τη μακροδομή του (Γιάγκου 2009: 154-155). Σε αυτά ακριβώς τα στοιχεία εστίασαν ανεξαιρέτως τόσο οι πρόδρομοι των ερευνών της αναγνωσιμότητας Sherman (1893) και Kitson (1921) όσο και οι μεταγενέστεροι εισηγητές των καθιερωμένων σήμερα δεικτών μέτρησής της από τα τέλη του 19ου αιώνα έως σήμερα (βλ. DuBay 2004: 13-18), χωρίς βέβαια να παραγνωρίζουν τη σημασία των υπόλοιπων παραγόντων της αναγνωσιμότητας.

1.2. Απαρχές και εφαρμογές

Οι πρώτες μελέτες που έθεσαν τα θεμέλια των ερευνών και των εφαρμογών αναγνωσιμότητας χρονολογούνται από τη δεκαετία του '20. Πρώτα ο Thorndike (1921) εστίασε στις πιο συχνόχρηστες λέξεις, καταρτίζοντας σχετικούς καταλόγους για διδακτική χρήση. Παρατήρησε ότι οι πιο συχνές λέξεις είναι συνάμα απλές στη σημασία και βραχείες και ανέδειξε τον καταλυτικό ρόλο που παίζει το λεξιλόγιο στην ανάπτυξη της αναγνωστικής ικανότητας του ατόμου. Συνεχιστής του υπήρξε ο Zipf (1949), ο οποίος διεξήγαγε στατιστικές αναλύσεις κειμένων για να δείξει πώς λειτουργεί η αρχή της ήσσονος προσπάθειας στη χρήση της γλώσσας. Το γενικό συμπέρασμα από τις εργασίες των δύο αυτών μελετητών ήταν ότι όσο πιο μικρές και απλές είναι οι λέξεις τόσο πιο συχνόχρηστες και εύκολες στη μάθηση/κατανόηση αποδεικνύονται. Ο Kitson (1921) προχώρησε λίγο ακόμη οδηγώντας την έρευνα πολύ κοντά στη δημιουργία των πρώτων *τύπων υπολογισμού αναγνωσιμότητας* (readability formulas, στο εξής TYA). Συγκεκριμένα, ήταν ο πρώτος που εντόπισε τη σημασία του μήκους των λέξεων και των προτάσεων ως μεταβλητών αναγνωσιμότητας. Αυτά τα δύο κειμενικά χαρακτηριστικά θα υιοθετηθούν αργότερα από τον Flesch (βλ. § 2.1), τον εισηγητή ενός από τους πιο διαδεδομένους TYA. Λίγα χρόνια αργότερα κατά τη δεκαετία του '30, σε μια περίοδο που η κυβέρνηση των Ηνωμένων Πολιτειών επενδύει στην εκπαίδευση ενηλίκων, ο Bryson θα διατυπώσει τη θέση ότι η κύρια αιτία της αναγνωστικής αποτυχίας είναι η έλλειψη αναγνωστικών δεξιοτήτων ως συνέπεια ελλιπούς εκπαίδευσης. Πίστευε, επίσης, ότι το άμεσο και διαυγές ύφος έρχεται ως αποτέλεσμα «μιας πειθαρχίας και μιας καλλιτεχνίας την οποία λίγοι άνθρωποι που έχουν ιδέες κάνουν τον κόπο να επιτύχουν. Εάν το απλό γράνιμο ήταν εύκολο, πολλά από τα προβλήματά μας θα είχαν επιλυθεί προ

πολλού» (βλ. Klare & Buck 1954: 58). Το έργο του είχε μεγάλη επίδραση στους μαθητές του Lorge και Flesch. Την ίδια περίπου περίοδο πολύ σημαντικές υπήρξαν οι μελέτες των Gray and Leary (1935), οι οποίοι στο έργο τους *“What makes a book readable”* εντόπισαν μια μακρά σειρά από στοιχεία τα οποία επηρεάζουν την αναγνωσιμότητα και τα ομαδοποίησαν σε τέσσερις κατηγορίες: α) περιεχόμενο, β) ύφος, γ) σχεδιασμός και δ) δομή. Πολύ σύντομα, ο Lorge (1939, 1944) με τις έρευνές του για τη σημασία του λεξικού σημασιολογικού φορτίου συνέβαλε αποφασιστικά στη μείωση του αριθμού των μεταβλητών που επηρεάζουν την αναγνωσιμότητα, προκρίνοντας τα εξής τρία μόνο χαρακτηριστικά του γραπτού λόγου: α) μέσο μήκος πρότασης σε λέξεις, β) αριθμός ονοματικών/ρηματικών φράσεων ανά 100 λέξεις και γ) αριθμός δύσκολων λέξεων που δεν συμπεριλαμβάνονται στον κατάλογο των 769 εύκολων λέξεων των Dale και Chall (1948). Με την εργασία του ο Lorge έθεσε τα θεμέλια για την εμφάνιση και διάδοση μετά από λίγα χρόνια των TYA των Dale and Chall (1948) και Flesch (1948 – 1975).

Πριν παρουσιάσουμε τους πιο σημαντικούς TYA, στους οποίους βασίζεται και η παρούσα έρευνα, θα αναφερθούμε εν συντομία στις κυριότερες εφαρμογές των μελετών αναγνωσιμότητας που καθιερώθηκαν από τα μέσα του 20^{ου} αιώνα και εξής (βλ. DuBay 2004: 54-55). Κατ’ αρχάς, δόθηκε ιδιαίτερη σημασία στις μετρήσεις αναγνωσιμότητας των δημοσιογραφικών εντύπων με σκοπό την αύξηση του αναγνωστικού κοινού τους. Επίσης, έγιναν μετρήσεις σε δικαστικές αποφάσεις και νομικά κείμενα σε μια προσπάθεια να προστατευθεί το δικαίωμα στη σαφή πληροφόρηση. Ακόμη, οι έρευνες αναγνωσιμότητας συνέβαλαν στη συγγραφή απλούστερου πληροφοριακού υλικού (ασφαλιστικά συμβόλαια, κείμενα με συμβουλές υγείας, οδηγίες χρήσης συσκευών, κ.λπ.) και στην ανάπτυξη διδακτικού υλικού κατάλληλα προσαρμοσμένου στις ανάγκες της εκάστοτε βαθμίδας εκπαίδευσης. Τέλος, υπήρξε μεγάλη η συμβολή των μελετών αναγνωσιμότητας γενικότερα στη γλωσσική έρευνα.

2. Τύποι υπολογισμού αναγνωσιμότητας

2.1. TYA Flesch

Ο Flesch, αυστριακός στην καταγωγή και ειδικός στη βιβλιοθηκονομία και την εκπαίδευση ενηλίκων, υπήρξε ο πρώτος που κατάφερε να προβάλει ευρύτερα τη σημασία των TYA. Έχοντας διεξαγάγει διάφορες έρευνες κυρίως σε αναγνωστικό υλικό για ενηλίκους, εισηγήθηκε έναν TYA (Flesch 1948), ο οποίος περιελάμβανε δύο μέρη: α) Τον δείκτη Reading Ease Score, που χρησιμοποιούσε δύο μεταβλητές, τον αριθμό των συλλαβών και τον αριθμό των προτάσεων ανά 100 λέξεις. Ο δείκτης αυτός προβλέπει την ευκολία της ανάγνωσης σε μια κλίμακα από το 1 έως το 100, όπου το 30 δηλώνει «πολύ δύσκολο» κείμενο και το 70 «εύκολο». β) Πρόβλεψη του ενδιαφέροντος του αναγνώστη μέσω μέτρησης του αριθμού των λέξεων με αναφορά σε πρόσωπα, όπως αντωνυμίες και ονόματα, καθώς και του αριθμού των προτάσεων «υποκειμενικού» χαρακτήρα,

όπως ο ευθύς λόγος και οι επιφωνηματικές και ελλειπτικές προτάσεις. Ο Reading Ease Score αναλύει ποσοτικά τις παραπάνω μεταβλητές κάθε κειμένου, αποδίδοντάς του έναν *δείκτη αναγνωσιμότητας* (readability index, στο εξής ΔΑ, βλ. Γιάγκου 2009: 158), ο οποίος αυξάνεται αντιστρόφως ανάλογα με την αναγνωστική δυσκολία του κειμένου, κατατάσσοντάς το σε ένα από τα επτά επίπεδα της εκατοντάβαθμης κλίμακας αναγνωστικής δυσκολίας του Flesch (βλ. πίνακα 1, σελ. 7). Ο TYA του Flesch είναι ο εξής (βλ. Γιάγκου 2009: 162):

$$R = 206,835 - (1,015 \times W/T) - (84,6 \times S/W)$$

Όπου: R = βαθμός στην κλίμακα αναγνωστικής δυσκολίας από το 0 (δύσκολο) έως το 100 (εύκολο)

W = αριθμός λέξεων

T = αριθμός προτάσεων

S = αριθμός συλλαβών

Αυτός ο TYA γνώρισε πολύ μεγάλη διάδοση λόγω της ικανοποιητικής αξιοπιστίας του¹, ασκώντας συνεπακόλουθα μεγάλη επιρροή στο δημοσιογραφικό λόγο. Τα αποτελέσματα μιας δεκάχρονης έρευνας που διεξήγαγε ο Flesch (1949), με μετρήσεις αναγνωσιμότητας στα κείμενα διαφόρων αμερικανικών περιοδικών, έδειξαν ότι, για παράδειγμα, τα κείμενα του *Reader's digest* έχουν ΔΑ 65, δηλαδή κατατάσσονται στο «μέσο» επίπεδο, ενώ το περιοδικό *American magazine* λαμβάνει ΔΑ μεταξύ 50 και 60, με αποτέλεσμα τα κείμενά του να αξιολογούνται ως ελαφρώς πιο δύσκολα από αυτά του προηγούμενου περιοδικού, δηλαδή στο επίπεδο «αρκετά δύσκολο».

Το 1975, στο πλαίσιο έρευνας χρηματοδοτούμενης από το αμερικανικό πολεμικό ναυτικό, τροποποιήθηκε ο TYA Reading Ease του Flesch, ώστε ο ΔΑ που θα προκύπτει να αντιστοιχίζει τα επτά επίπεδα της αξιολογικής κλίμακας του TYA Flesch με τις διάφορες βαθμίδες εκπαίδευσης των ΗΠΑ (Kincaid et al. 1975). Ο νέος, λοιπόν, TYA γνωστός ως Flesch-Kincaid, ή απλώς Kincaid, είναι ο εξής (βλ. Γιάγκου 2009: 162):

$$G = 0,39 \times W/T + 11,8 \times S/W - 15,59$$

Όπου: G = τάξη (του εκπαιδευτικού συστήματος των ΗΠΑ) για την οποία είναι κατάλληλο το κείμενο

W = αριθμός λέξεων

T = αριθμός προτάσεων

S = αριθμός συλλαβών

1 Ενδεικτικά, η τιμή συσχέτισης του TYA Flesch με τα τεστ αναγνωστικής κατανόησης των McCall και Crabbs (1950) είναι .64, η οποία δηλώνει ένα ικανοποιητικό επίπεδο αξιοπιστίας.

Η αντιστοίχιση μεταξύ της αξιολογικής κλίμακας του Flesch Reading Ease και των εκπαιδευτικών βαθμίδων του TYA Kincaid, όπως προσαρμόστηκαν στα ελληνικά από τους Ευσταθιάδη κ.ά. (2002), είναι η εξής (βλ. και DuBay 2004: 22):

Πίνακας 1.

Αξιολογική κλίμακα Flesch-Kincaid προσαρμοσμένη στο ελληνικό εκπαιδευτικό σύστημα

Flesch Reading Ease		Flesch-Kincaid Grade Level
100 – 90	πολύ εύκολο	Α΄-Β΄-Γ΄ Δημοτικού
90 – 80	εύκολο	Δ΄-Ε΄-ΣΤ΄ Δημοτικού
80 – 70	αρκετά εύκολο	Α΄-Β΄ Γυμνασίου
70 – 60	μέσο	Γ΄ Γυμν. Α΄ Λυκείου
60 – 50	αρκετά δύσκολο	Β΄- Γ΄ Λυκείου
50 – 30	δύσκολο	ΑΕΙ – ΤΕΙ
30 – 0	πολύ δύσκολο	ΑΕΙ – ΤΕΙ

Ενδεικτικό της ευρείας διάδοσης του συγκεκριμένου TYA είναι το γεγονός ότι εφαρμόστηκε και σε άλλες γλώσσες εκτός των αγγλικών, όπως τα γαλλικά, τα γερμανικά, τα ολλανδικά και τα ιταλικά.

2.2. TYA Fog

Λίγο μετά τις πρωτοποριακές έρευνες του Flesch, ο Robert Gunning (1952) εισηγήθηκε τον TYA Fog, ο οποίος έγινε δημοφιλής χάρη στην πολύ απλή μορφή του. Ο συγκεκριμένος TYA χρησιμοποιεί δύο μεταβλητές: α) το μέσο μήκος κάθε πρότασης και β) τον αριθμό των λέξεων άνω των δύο συλλαβών ανά 100 λέξεις και είναι ο εξής (βλ. Γιάγκου 2009: 162):

$$G = 0,4 + W/T + 100 \times P_2/W$$

Όπου: G = τάξη (του εκπαιδευτικού συστήματος των ΗΠΑ) για την οποία είναι κατάλληλο το κείμενο

W = αριθμός λέξεων

T = αριθμός προτάσεων

P₂ = αριθμός λέξεων με δύο ή περισσότερες συλλαβές

Με τις έρευνές του ο Gunning καθιέρωσε τη χρήση δεικτών αναγνωσιμότητας σε πολλές εφημερίδες και περιοδικά των ΗΠΑ με σκοπό την αποφυγή της «περιττής στρυφνότητας», η οποία, σύμφωνα με τους ισχυρισμούς του, αποτελεί την κύρια αιτία της αναγνωστικής αδυναμίας των μαθητών.

2.3. TYA SMOG

Ένας άλλος ειδικός της αναγνωσιμότητας, ο Harry G. McLaughlin (1969), δημοσίευσε τον TYA SMOG με την πεποίθηση ότι το μήκος των λέξεων πρέπει να πολλαπλασιάζεται με το μήκος των προτάσεων, παρά να προστίθεται σ' αυτό. Ως εκ τούτου, βασική μεταβλητή αυτού του TYA είναι ο αριθμός των λέξεων με τρεις ή περισσότερες συλλαβές ανά 30 προτάσεις (βλ. DuBay 2004: 47 και Γιάγκου 2009: 162):

$$G = 3,1291 + 1,043 \times \sqrt{P_3}$$

Όπου: G = τάξη (του εκπαιδευτικού συστήματος των ΗΠΑ) για την οποία είναι κατάλληλο το κείμενο

P_3 = αριθμός λέξεων με τρεις ή περισσότερες συλλαβές

Σημειώτεον ότι ο ΔΑ τόσο του TYA Fog όσο και του TYA SMOG αυξάνει ανάλογα με το επίπεδο αναγνωστικής δυσκολίας ενός κειμένου.

2.4. Ασυμφωνίες των τύπων υπολογισμού αναγνωσιμότητας

Σε αυτό το σημείο είναι χρήσιμο να επισημανθούν οι ασυμφωνίες που παρατηρούνται μεταξύ των διαφόρων TYA και τις οποίες κάθε ερευνητής οφείλει να έχει υπόψη του, εάν σκοπεύει να τους αξιοποιήσει ερευνητικά. Κατ' αρχάς, όπως κατέστη φανερό από την προηγούμενη παρουσίαση των τεσσάρων TYA Flesch, Flesch-Kincaid, Fog και SMOG, οι TYA χρησιμοποιούν διαφορετικές μεταβλητές μέτρησης και επομένως διαφορετικούς αλγόριθμους για τη μέτρηση της αναγνωσιμότητας. Επίσης, χρησιμοποιούν διαφορετική *βαθμολογική βάση* (criterion score). Η έννοια της βαθμολογικής βάσης σχετίζεται με τις επικρατέστερες δοκιμές αναγνωστικής κατανόησης των μαθητών που χρησιμοποιούνται για την ανάπτυξη και την επαλήθευση των TYA στις Ηνωμένες Πολιτείες. Οι ερευνητές, δηλαδή, προκειμένου να επαληθεύσουν την αξιοπιστία του TYA που προτείνουν, χρησιμοποιούν τις αξιολογικές κλίμακες-επίπεδα αυτών των δοκιμών αναγνωστικής κατανόησης ως σημείο αναφοράς για την ακρίβεια των μετρήσεών τους. Μία από τις πιο αναγνωρισμένες δοκιμές αυτού του είδους είναι των McCall και Crabbs (1926, 1950, 1961, 1979) με τον τίτλο *Standard Test Lessons in Reading*, η οποία περιέχει ασκήσεις κατανόησης γραπτού λόγου με ερωτήσεις πολλαπλής επιλογής σε κείμενα διαβαθμισμένης δυσκολίας. Ακριβώς, όμως, επειδή οι TYA που συσχετίζονται με αυτήν την αξιολογική κλίμακα δεν είναι ίδιοι ως προς τα διάφορα χαρακτηριστικά τους, κατά την ανάπτυξή τους δε χρησιμοποιείται πάντοτε η ίδια βαθμολογική βάση για την κατάταξη ενός κειμένου σε κάποιο επίπεδο αναγνωσιμότητας. Για παράδειγμα, ο δείκτης Fog αναπτύχθηκε με βάση κατάταξης το 90% των σωστών απαντήσεων σε δοκιμή αναγνωστικής κατανόησης, ενώ ο δείκτης SMOG με το 100% των σωστών απαντήσεων. Με δεδομένες τις προηγούμενες ασυμφωνίες, το στοιχείο που προσδίδει

εν τέλει εγκυρότητα σε έναν δείκτη αναγνωσιμότητας είναι ο βαθμός της συνέπειάς του ως προς την πρόβλεψη της δυσκολίας ανάγνωσης μιας σειράς κειμένων διαβαθμισμένης δυσκολίας (DuBay 2004: 56).

2.5. Περιορισμοί των ΤΥΑ

Οι ιδιαιτερότητες των ΤΥΑ που περιγράφηκαν στην προηγούμενη παράγραφο μάς οδηγούν στη διατύπωση των περιορισμών τους. Ουσιαστικά, αυτοί οι ΤΥΑ δεν παρέχουν την τέλεια πρόβλεψη, αλλά μια κατά προσέγγιση εκτίμηση της αναγνωστικής δυσκολίας ενός κειμένου, γεγονός που αποδεικνύεται και από τους διαφορετικούς ΔΑ που δίνουν οι διάφοροι ΤΥΑ για το ίδιο κείμενο (DuBay 2004: 56). Για αυτό το λόγο κατά την αποτίμηση της δυσκολίας ενός κειμένου πρέπει να συνεκτιμώνται και άλλοι ρητορικοί παράγοντες, όπως η δομή, το περιεχόμενο, η συνεκτικότητα και η μορφοποίηση του κάθε κειμένου. Εξάλλου, οι ΤΥΑ έχουν κατά καιρούς δεχθεί αυστηρή κριτική, η οποία εστιάζει αφενός μεν στο γεγονός ότι δεν λαμβάνονται καθόλου υπόψη παράμετροι που αφορούν τον ίδιο τον αναγνώστη και αφετέρου στο γεγονός της αναξιοπιστίας των ίδιων των επιφανειακών χαρακτηριστικών που αξιοποιούν οι ΤΥΑ ως δεικτών ευκολίας ή δυσκολίας ενός κειμένου².

2.6. Βασικές παραδοχές για τους ΤΥΑ

Έχοντας παρουσιάσει κάποια βασικά χαρακτηριστικά καθώς και τους περιορισμούς των ΤΥΑ, είναι σκόπιμο να σημειωθούν συνοπτικά ορισμένες βασικές παραδοχές για αυτούς, οι οποίες θα φωτίσουν καλύτερα τη φύση τους (βλ. DuBay 2004: 19-20).

Όσον αφορά τον αριθμό των μεταβλητών που χρησιμοποιεί ένας ΤΥΑ, το αν είναι μικρός ή μεγάλος δεν έχει ιδιαίτερο αντίκτυπο στην ακρίβεια των προβλέψεων, αλλά σίγουρα όσο περισσότερες μεταβλητές λαμβάνει υπόψη του ένας ΤΥΑ τόσο πιο δύσκολος θεωρείται. Συνεπώς, η έως τώρα έρευνα και οι εφαρμογές της αναγνωσιμότητας έχουν δείξει ότι δύο είναι οι μεταβλητές με τη βέλτιστη δυνατότητα πρόβλεψης της αναγνωστικής δυσκολίας: α) η σημασιολογική, η οποία αφορά μετρήσεις δυσκολίας του λεξιλογίου, από τις οποίες η πιο επιτυχημένη είναι η μέτρηση του αριθμού των μορφημάτων ανά 100 λέξεις και β) η συντακτική, όπου έχουμε μέτρηση του μέσου μήκους των προτάσεων. Πιο συγκεκριμένα, έχει αποδειχθεί ότι αφενός ο μέσος αριθμός των συλλαβών ανά 100 λέξεις (Coleman 1971) και αφετέρου ο μέσος αριθμός λέξεων ανά πρόταση (Bormuth 1966), αποτελούν πολύ πιο εύκολους τρόπους πρόβλεψης της δυσκολίας ενός κειμένου και σχεδόν το ίδιο ακριβείς με τη μέτρηση των μορφημάτων.

² Για εκτενέστερη αναφορά στις κριτικές αυτές βλ. Γιάγκου 2009: 164 κ. εξ.

2.7. Ο TYA Flesch για την ελληνική γλώσσα

Στις αρχές της δεκαετίας του 1980 ο Γαγάτσης ήταν ο πρώτος που ασχολήθηκε με την αναγνωσιμότητα ελληνικών κειμένων, μελετώντας την αναγνωσιμότητα των σχολικών εγχειριδίων των Μαθηματικών και προσαρμοσε τον TYA Flesch Reading Ease Score στα ελληνικά (Γαγάτσης 1985). Στηριζόμενος στην παρατήρηση ότι οι ελληνικές λέξεις είναι κατά μέσο όρο μεγαλύτερες από τις αγγλικές ή τις γαλλικές, αντικατέστησε το συντελεστή 84,6 του μέσου αριθμού συλλαβών ανά λέξη με τον 59. Έτσι ο TYA Flesch Reading Ease Score προσαρμοσμένος στα ελληνικά αναδιαμορφώνεται ως εξής:

$$R^3 = 206,835 - (1,015 \times W/T) - (59 \times S/W)$$

3. Λογισμικό αναγνωσιμότητας ελληνικών κειμένων

Στις αρχές της πρώτης δεκαετίας του 21^{ου} αιώνα δημιουργείται από το Κέντρο Ελληνικής Γλώσσας (ΚΕΓ) το gival 1.1 (από το Greek evaluation), ένα νέο λογισμικό για την αποτίμηση της αναγνωσιμότητας ελληνικών κειμένων (Ευσταθιάδης κ.ά. 2002), το οποίο στηρίζεται σε προσαρμογές στα Νέα Ελληνικά των TYA Flesch Reading Ease (μέσο μήκος προτάσεων, μέσος αριθμός συλλαβών), Flesch-Kincaid Grade Level (επίπεδα εκπαίδευσης), Flesch Fog Index (προθήματα, επιθήματα) και Smog (πολυσύλλαβες λέξεις). Το συγκεκριμένο λογισμικό είναι διαθέσιμο για το κοινό μέσω της ιστοσελίδας της «Πύλης για την Ελληνική Γλώσσα» του ΚΕΓ³. Όπως επισημαίνει ο Ευσταθιάδης κ.ά. (2002), κατά την ανάπτυξή του το λογισμικό δοκιμάστηκε επανειλημμένως στο κειμενικό υλικό των εξεταστικών θεμάτων ελληνομάθειας του ΚΕΓ των ετών 1999, 2000 και 2001, και σε κάθε περίπτωση ανάλογα με τα αποτελέσματα που εξάγονταν επιφέρονταν σταδιακές βελτιώσεις σε αυτό, προκειμένου να επιτευχθεί η καλύτερη δυνατή προσαρμογή του στην ελληνική γλώσσα. Ο βασικός στόχος των σχεδιαστών του λογισμικού ήταν να δίνει μία γενική άποψη της δυσκολίας ανάγνωσης ενός κειμένου με βάση αντικειμενικές ενδείξεις από τις μετρήσεις των διαφόρων επιφανειακών κειμενικών χαρακτηριστικών. Επίσης, ο κύριος σκοπός της δημιουργίας του ήταν να συμβάλλει σε μια πρώτη επιλογή των κειμένων για τις δεξιότητες κατανόησης προφορικού και γραπτού λόγου των εξετάσεων πιστοποίησης επάρκειας της ελληνομάθειας του ΚΕΓ καθώς και στη διαβάθμισή τους τόσο μεταξύ των επιπέδων ελληνομάθειας όσο και εσωτερικά σε κάθε επίπεδο.

Στη συνέχεια παρατίθενται ενδεικτικά οι βασικές διεπαφές του λογισμικού αναγνωσιμότητας ελληνικών κειμένων gival 1.1, όπως εμφανίζονται στην ιστοσελίδα της Πύλης για την Ελληνική Γλώσσα:

3 Για επεξηγήσεις των συντομεύσεων του τύπου υπολογισμού της αναγνωσιμότητας βλ. § 2.1.

4 Ο ακριβής σύνδεσμος του λογισμικού είναι: http://www.greek-language.gr/greekLang/modern_greek/foreign/tools/readability/index.html.

Εικόνα 1. Αρχική διεπαφή του gval 1.1 για την εισαγωγή κειμένου

Στην εικόνα 1 διακρίνεται το κενό πεδίο στο οποίο ο χρήστης μπορεί μέσω των εντολών «αντιγραφή» και «επικόλληση» του ηλεκτρονικού υπολογιστή να εισαγάγει το προς αξιολόγηση κείμενο με τον τρόπο που φαίνεται στην εικόνα 2:

Εικόνα 2. Εισαγωγή του προς αξιολόγηση κειμένου στο λογισμικό gval 1.1

Σε αυτό το σημείο ο χρήστης επιλέγοντας την εντολή «Αποστολή» λαμβάνει τους ΔΑ του κειμένου με την εξής μορφή:

Εικόνα 3. Αποτελέσματα της μέτρησης αναγνωσιμότητας του λογισμικού grval 1.1

Όπως φαίνεται στην εικόνα 3, το λογισμικό grval 1.1 δίνει μετρήσεις για τον αριθμό των προτάσεων (T), των λέξεων (W), των συλλαβών (S), για το μέσο αριθμό των λέξεων ανά πρόταση (W/T) και το μέσο αριθμό συλλαβών ανά λέξη (S/W). Τέλος, εμφανίζονται οι ΔΑ από τους τέσσερις TYA που χρησιμοποιούνται. Ο πρώτος ΔΑ είναι σύμφωνα με τους TYA Flesch και Kincaid, ο δεύτερος ΔΑ είναι του SMOG και ο τρίτος είναι του Fog.

4. Στόχοι και υποθέσεις έρευνας

Κύριος στόχος της παρούσας έρευνας είναι η περαιτέρω διερεύνηση της αξιοπιστίας του λογισμικού grval 1.1, βάσει στοιχείων από την εφαρμογή του σε κείμενα πιο πρόσφατων εξεταστικών θεμάτων και συνεπακόλουθα η εξεύρεση τρόπων για την περαιτέρω βελτίωσή του στην περίπτωση που τα εμπειρικά δεδομένα δείξουν μη ικανοποιητικό επίπεδο συνέπειας.

Η βασική υπόθεση από την οποία εκκινεί η έρευνα είναι ότι η δυσκολία κατανόησης των κειμένων των εξεταστικών θεμάτων κατανόησης γραπτού λόγου για την πιστοποίηση της ελληνομάθειας αυξάνει βαθμιαία από τα κατώτερα στα ανώτερα επίπεδα ελληνομάθειας, παρουσιάζοντας αισθητή διαφορά από επίπεδο σε επίπεδο (βλ. Ευσταθιάδης κ.ά. 2002). Το κρίσιμο ζήτημα είναι, επομένως, εάν η παραπάνω διαφορά ανιχνεύεται σε ικανοποιητικό βαθμό από το «Λογισμικό Αναγνωσιμότητας Ελληνικών Κειμένων grval 1.1».

5. Μεθοδολογία

5.1. Γλωσσικό υλικό

Ως υλικό έρευνας χρησιμοποιήθηκαν όλα τα κείμενα από τα εξεταστικά θέματα της πιστοποίησης επάρκειας της ελληνομάθειας (ΠΕΕ) του Κέντρου Ελληνικής Γλώσσας για τη δεξιότητα της κατανόησης γραπτού λόγου σε όλα τα επίπεδα (Α, Β, Γ, Δ) των ετών 2007, 2008, 2009 και 2010. Συγκεκριμένα, από κάθε έτος συγκεντρώθηκαν τα δύο κείμενα 200-250 λέξεων του επιπέδου Α, τα δύο κείμενα 250-300 λέξεων του επιπέδου Β, τα δύο κείμενα 400-500 λέξεων του επιπέδου Γ και τα τρία κείμενα 600-800 λέξεων του επιπέδου Δ.⁵ Συνολικά, το πλήθος των κειμένων που ερευνήθηκαν ήταν 36.

5.2. Επεξεργασία γλωσσικού υλικού

Η επεξεργασία του γλωσσικού υλικού που συλλέχθηκε διεξήχθη στις εξής φάσεις:

- α) εισαγωγή των κειμένων στο gval 1.1 και εξαγωγή ΔΑ για κάθε ένα από αυτά,
- β) καταχώριση των ΔΑ σε λογιστικό φύλλο και
- γ) στατιστική επεξεργασία (t-test) των ΔΑ.

Κατά την τελευταία φάση, έγιναν συγκριτικές μετρήσεις ως προς τους διάφορους ΔΑ που έδωσαν οι ΤΥΑ του gval 1.1 μεταξύ: α) όλων των κειμένων κατανόησης γραπτού λόγου και β) των επιπέδων ελληνομάθειας ανά τα εξής τρία ζεύγη: 1) Α – Β, 2) Β – Γ και 3) Γ – Δ.

6. Αποτελέσματα

Στο γράφημα 1 εμφανίζονται οι μέσες τιμές των ΔΑ που έδωσαν οι ΤΥΑ για όλα τα κείμενα από τα εξεταστικά θέματα της κατανόησης γραπτού λόγου και των τεσσάρων ετών του ερευνητικού δείγματος ανά επίπεδο ελληνομάθειας και ανά μέρος δεξιότητας. Στα επίπεδα Α, Β και Γ η δεξιότητα κατανόησης γραπτού λόγου έχει δύο μέρη, ενώ στο επίπεδο Δ τρία μέρη.

Στη συνέχεια παρουσιάζονται σε πίνακες οι μέσες τιμές των ΔΑ ανά ΤΥΑ και ανά επίπεδο. Στον πίνακα 2 σημειώνονται τα αποτελέσματα για τους ΤΥΑ Flesch και Kincaid:

Μία πρώτη παρατήρηση στα παραπάνω αποτελέσματα είναι ότι ο μέσος όρος των ΔΑ παρουσιάζει στατιστικά σημαντική διαφορά μόνο μεταξύ των επιπέδων Β και Γ ($p=0,018$).

Στον πίνακα 3 εμφανίζονται τα αποτελέσματα για τον ΤΥΑ SMOG:

⁵ Σημειωτέον ότι οι συγκεκριμένες ονομασίες των επιπέδων έχουν μεταβληθεί από το 2010 ως εξής: Α → Α2, Β → Β1, Γ → Β2, Δ → Γ1.

Γράφημα 1. Μέσες τιμές των ΔΑ του grval 1.1 για τα κείμενα των ετών 2007 έως 2010

Πίνακας 2.

Μέσες τιμές των ΔΑ του TYA Flesch ανά επίπεδο ελληνομάθειας

Επίπεδο ελληνομάθειας	Μέσος όρος των ΔΑ	Βαθμίδα Flesch	Βαθμίδα Kincaid
A	81,07 (SD=7,510)	εύκολο	Δ-Ε-ΣΤ Δημοτ.
B	79,30 (SD=8,725)	αρκετά εύκολο	A-B Γυμν.
Γ	65,90 (SD=12,443)	μέσο	Γ Γυμν. – Α Λυκ.
Δ	59,48 (SD=8,840)	αρκετά δύσκολο	B-Γ Λυκ.

Πίνακας 3.

Μέσες τιμές των ΔΑ του TYA SMOG ανά επίπεδο ελληνομάθειας

Επίπεδο ελληνομάθειας	Μέσος όρος των ΔΑ	Βαθμίδα SMOG
A	8,19 (SD=0,94)	A' Γυμν.
B	8,71 (SD=1,57)	A' Γυμν.
Γ	10,84 (SD=1,78)	Γ' Γυμν.
Δ	11,96 (SD=1,95)	A' Λυκ.

Εδώ παρατηρείται ότι ο μέσος όρος των ΔΑ των κειμένων είναι σημαντικά διαφορετικός τόσο μεταξύ των επιπέδων B και Γ ($p=0,007$) όσο και μεταξύ των επιπέδων Γ και Δ ($p=0,017$).

Τέλος, στον πίνακα 4 παρουσιάζονται τα αποτελέσματα για τον TYA Fog:

Πίνακας 4.

Μέσες τιμές των ΔΑ του TYA Fog ανά επίπεδο ελληνομάθειας

Επίπεδο ελληνομάθειας	Μέσος όρος των ΔΑ	Βαθμίδα FOG
A	10,13 (SD=4,33)	πολύ εύκολο
B	12,01 (SD=7,2)	πολύ εύκολο
Γ	18,83 (SD=9,51)	εύκολο
Δ	24,87 (SD=7,24)	αρκετά εύκολο

Και σ' αυτή την περίπτωση, δεν υπήρξαν μεγάλες διαφορές μεταξύ των διαδοχικών επιπέδων, καθώς ο μέσος όρος των ΔΑ εμφανίζει στατιστικά σημαντικές διαφορές μόνο μεταξύ των επιπέδων Γ και Δ ($p=0,008$).

7. Συζήτηση - συμπεράσματα

Με βάση τα αποτελέσματα που παρουσιάστηκαν προηγουμένως, μπορούμε να εξαγάγουμε ορισμένα συμπεράσματα. Κατ' αρχάς, όλοι οι TYA διαβαθμίζουν σε γενικές γραμμές τα κείμενα όλων των επιπέδων ελληνομάθειας με την αναμενόμενη σειρά δυσκολίας. Με αυτό εννοούμε ότι, με εξαίρεση ίσως τα επίπεδα Α και Β των οποίων τα κείμενα δεν παρουσίασαν μεταξύ τους στατιστικά σημαντική διαφορά ως προς τους ΔΑ, δεν παρατηρήθηκε κάποια συστηματική αξιολόγηση από κανέναν από τους τέσσερις TYA κειμένων κατώτερων επιπέδων ως πιο δύσκολων από κείμενα των ανώτερων επιπέδων. Αντίθετα, σύμφωνα με τα παραπάνω αποτελέσματα, οι αναμενόμενες διαφορές μεταξύ των ΔΑ που έδωσαν οι TYA Flesch-Kincaid και SMOG για τα επίπεδα Β και Γ ήταν στατιστικά σημαντικές. Το ίδιο συνέβη μεταξύ των κειμένων των επιπέδων Γ και Δ όσον αφορά τις αξιολογήσεις των TYA SMOG και Fog. Παράλληλα, όμως, με αυτή τη γενικότερη τάση για αναμενόμενους ΔΑ υπήρξαν και περιπτώσεις κειμένων, λίγων είναι η αλήθεια, τα οποία δεν κατατάχθηκαν από τους δείκτες στις αναμενόμενες βαθμίδες εκπαίδευσης. Το φαινόμενο αυτό παρατηρήθηκε ιδιαίτερα στις αξιολογήσεις των TYA SMOG και Fog. Συγκεκριμένα, υπήρξαν αφενός κείμενα από εξεταστικά θέματα για το ίδιο επίπεδο ελληνομάθειας, των οποίων οι ΔΑ είχαν πολύ μεγάλη διαφορά μεταξύ τους και αφετέρου κείμενα από διαφορετικά επίπεδα με πολύ μικρές διαφορές. Για παράδειγμα, υπήρξε κείμενο του επιπέδου Δ που έλαβε από τον Flesch-Kincaid ΔΑ 72,21, ο οποίος το κατέταξε στη βαθμίδα «αρκετά εύκολο» του Flesch, που αντιστοιχεί στην βαθμίδα «Α'-Β' Γυμνασίου» της κλίμακας Kincaid, η οποία με τη σειρά της αντιστοιχεί κατά προσέγγιση στο επίπεδο Β ή Γ (Σακελλαρίου & Ευσταθιάδης 2002). Πρόκειται για ΔΑ που αποκλίνει αρκετά από τους ΔΑ που έλαβαν τα περισσότερα κείμενα του συγκεκριμένου επιπέδου, οι οποίοι κυμαίνονταν μεταξύ του 60 και του 48. Όσον αφορά

την άλλη περίπτωση, δύο κείμενα από το επίπεδο Α έλαβαν από τον ίδιο ΤΥΑ τιμές ΔΑ που κινήθηκαν μεταξύ του 73 και του 68. Αυτές, όμως, οι τιμές τα αξιολόγησαν ως ελάχιστα πιο δύσκολα από τέσσερα κείμενα του επιπέδου Γ, τα οποία έλαβαν ΔΑ μεταξύ 79 και 72.

Τα παραπάνω δεδομένα, ενισχύουν τη θέση ότι οι ΤΥΑ, με τη μορφή που έχουν στο λογισμικό gval 1.1, δεν μπορούν να κάνουν τίποτε παραπάνω από το να παρέχουν μια κατά προσέγγιση αξιολόγηση της αναγνωστικής δυσκολίας των ελληνικών κειμένων.

Με σκοπό τη βελτίωση της ακρίβειας των ΤΥΑ, η έρευνα θα μπορούσε να λάβει τις εξής κατευθύνσεις: α) προσθήκη του παράγοντα «εύκολες – δύσκολες σημασίες» μέσω της μέτρησης του ποσοστού των λέξεων σε κάθε κείμενο, οι οποίες δεν συμπεριλαμβάνονται στον κατάλογο των εύκολων/συχνόχρηστων λέξεων (πρβλ. ΤΥΑ Dale and Chall 1948), β) χρήση του παραπάνω παράγοντα συμπληρωματικά με τους υπόλοιπους παράγοντες του gval 1.1, γ) μέτρηση συντακτικών και γραμματικών παραμέτρων (μέρη του λόγου) που συμβάλλουν στη δυσκολία/ευκολία ενός κειμένου. Στην τελευταία περίπτωση, βέβαια, προϋποτίθεται απαραίτητως η χρήση ηλεκτρονικών εργαλείων, όπως μορφολογικού και συντακτικού αναλυτή για φυσική γλώσσα (βλ. Γιάγκου 2009 και Giagkou forthcoming).

Βιβλιογραφία

Bormuth, J.R. 1966. Readability: a new approach. *Reading Research Quarterly* 1: 79-132.

Γαγάτσης, Α. 1985. Η αναγνωσιμότητα των σχολικών βιβλίων των μαθηματικών του Δημοτικού Σχολείου. *Σύγχρονη Εκπαίδευση* 20: 40-48.

Γιάγκου, Μ. 2009. *Σώματα Κειμένων και Γλωσσική Εκπαίδευση: Δυνατότητες Αξιοποίησης στη Διδασκαλία της Ελληνικής και Συγκρότηση Παιδαγωγικά Κατάλληλων Σωμάτων Κειμένων*. Αδημοσίευτη διδακτορική διατριβή, Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών.

Coleman, E.B. 1971. Developing a technology of written instruction: some determiners of the complexity of prose. In E.Z. Rothkopf and P.E. Johnson (eds), *Verbal Learning Research and the Technology of Written Instruction*. New York: Teachers College Press, Columbia University, 155-204.

Dale, E. and J.S. Chall. 1948. A formula for predicting readability. *Educational Research Bulletin* 27(1): 11-20.

DuBay, W.H. 2004. *The Principles of Readability*. Costa Mesa, California: Impact Information.

Ευσταθιάδης, Σ., Ν. Αντωνοπούλου, Κ. Ζούρου και Θ. Τριανταφυλλίδης. 2002. Λογισμικό Αναγνωσιμότητας Ελληνικών Κειμένων. Αντλήθηκε στις 20 Σεπτεμβρίου 2011 από: http://www.greek-language.gr/greekLang/modern_greek/foreign/tools/readability/01.html

Flesch, R. 1948. A new readability yardstick. *Journal of Applied Psychology* 32: 221-233.

Flesch, R. 1949. *The Art of Readable Writing*. New York: Harper.

Giagkou, M. (forthcoming). Quantitative investigation of linguistic features as indices of Greek text difficulty. Poster presentation at the 10th International Conference of Greek Linguistics, September 1-4 2011, Komotini, Greece.

Gray, W.S. and B. Leary. 1935. *What Makes a Book Readable*. Chicago: Chicago University Press.

Gunning, R. 1952. *The Technique of Clear Writing*. New York: McGraw-Hill.

Kincaid, J.P., R.P. Fishburne, R.L. Rogers, and B.S. Chissom. 1975. Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel. *CNTECHTRA Research Branch Report* 8-75.

Kitson, H.D. 1921. *The Mind of the Buyer*. New York: Macmillan.

Klare, G.R. and B. Buck. 1954. *Know Your Reader, the Scientific Approach to Readability*. New York: Heritage House.

Lorge, I. 1939. Predicting reading difficulty of selections for children. *Elementary English Review* 16: 229-233.

Lorge, I. 1944. Predicting readability. *Teachers College Record* 45: 404-419.

McCall, W.A. and L.M. Crabbs. 1926, 1950, 1961, 1979. *Standard Test Lessons in Reading*. New York: Teachers College, Columbia University Press.

McLaughlin, H.G. 1969. SMOG grading – A new readability formula. *Journal of reading* 22: 639-646.

Σακελλαρίου, Α. και Σ. Ευσταθιάδης. 2002. Σύγκριση Προγραμμάτων Σπουδών. Σύγκριση του Αναλυτικού Εξεταστικού Προγράμματος Πιστοποίησης Επάρκειας της Ελληνομάθειας του Κέντρου Ελληνικής Γλώσσας προς το Πρόγραμμα Σπουδών για τη Γλώσσα όλων των Εκπαιδευτικών Βαθμίδων. Θεσσαλονίκη: Κέντρο Ελληνικής Γλώσσας. Αντλήθηκε στις 20 Σεπτεμβρίου 2011 από http://www.greeklanguage.gr/certification/sites/greeklanguage.gr.certification/files/Sigkrisi_programmatwn_

Sherman, A.L. 1893. *Analytics of Literature: A Manual for the Objective Study of English Prose and Poetry*. Boston: Ginn & Co.

Thorndike, E.L. 1921. *The Teacher's Word Book*. New York: Bureau of Publications, Teachers College, Columbia University.

Zipf, G.K. 1949. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. New York: Addison-Wesley.